DOCUMENT RESUME

ED 302 452 SO 019 519

AUTHOR Tohidi, Nayereh E.; And Others
TITLE Gender Differences in Performance on Tests of Cognitive Functioning: A Meta-Analysis of Research Findings.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 86
GRANT NSF-RISE-80-07857
NOTE 60p.; Paper based on a doctoral dissertation, University of Illinois.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Analysis of Variance; *Cognitive Measurement; College Students; Elementary School Students; Literature Reviews; Meta Analysis; Regression (Statistics); Secondary School Students; *Sex Differences; Social Science Research; *Statistical Surveys

ABSTRACT

An analysis of the findings in 70 U.S. and foreign studies that compare male and female students' performance on tests of cognitive functioning was conducted. The studies were reported in various science and psychology journals between 1965 and 1981. Overall, the results showed small but consistent gender differences, with males scoring slightly higher than females in cognitive domains involving conservation, spatial visualization, cognitive style, reasoning, and logical thinking; females performed slightly better than males on specific tasks involving classification and seriation. Analysis of variance and regression analyses revealed that the following variables had significant effects on the magnitude of gender differences: (1) cognitive domain; (2) response mode; (3) region of country; (4) sample selection; (5) type of community; (6) and year of publication. A general reference list of 85 items as well as tables and graphs are included. The appendices contain a list of 70 synthesized studies and a breakdown of the studies by cognitive domains. (DJC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302452

Gender Differences in Performance on Tests of Cognitive Functioning: A Meta-Analysis of Research Findings[1,2]

Nayereh E. Tohidi, Ph.D.

California State Polytechnic University, Pomona

Marjorie W. Steinkamp, Ph.D.

and

Martin L. Maehr, Ph.D.

University of Illinois at Urbana-Champaign

1986

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NAYEREH E. TOHIDI

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

30 819 519

BEST COPY AVAILABLE

Gender Differences in Performance on Tests of Cognitive Functioning: A Meta-Analysis of Research Findings

A meta-analysis of findings from 70 American and foreign studies reported in the literature between 1965 and 1981, comparing school-aged boys' and girls' performances on tests of cognitive functioning was conducted. Overall, the results showed small but consistent gender differences, with boys scoring slightly higher than girls in cognitive domains involving conservation, spatial visualization, cognitive style, reasoning, and logical thinking; girls performed slightly better than boys on specific tasks involving classification and seriation. Analysis of variance and regression analyses revealed that the following variables had significant effects on the magnitude of gender differences: cognitive domain, response mode, region of country, sample selection, type of community, and year of publication.

Introduction

A major problem continues to engage those concerned with fostering an equitable society: the gross underrepresentation of women in the higher echelons of technical and scientific professions (Astin, 1969; Cole, 1979). Moreover, the situation is not likely to change soon. Gender-related differences in academic preference and performance in disciplines essential to future scientific achievement are well documented (Armstrong, 1979; Steele & Wise, 1979). Even after artifacts of measurement, test content, and statistical analysis are taken into account, males' scores on standardized tests of mathematics and science are not only slightly higher than those of females, but they tend to persist in the face of cultural diversity (Husen, 1967; Keeves, 1973; Steinkamp, Harnisch, Walberg & Tsai, 1985). Although the existence of these differences is no longer in dispute, the reasons for such disparity are poorly understood. Research which sheds light on these issues, then, is clearly of highest priority.

Recently, considerable attention has been focused on the hypothesis that observed gender differences in representation and achievement in the world of science and mathematics originate in disparities in cognitive abilities (Benbow & Stanley, 1980). Moreover, several partial reviews of the literature support the notion of a male advantage on tests of cognitive ability. Schonberger (1976) reviewed a number of studies on problem-

solving and concluded that small gender-related differences start at early adolescence and increase until maturity. Steinkamp and Maehr (1983) searched the literature and on the basis of 68 gender comparisons on tests of cognitive ability reported that boys scored slightly but significantly higher than girls. Maccoby and Jacklin (1974) reviewed 30 studies of spatial ability and reported that a male advantage emerges in early adolescence and is maintained into adulthood. In her critical re-review of Maccoby and Jacklin's work, Sherman (1978) affirmed a slight but significant male advantage. When quantitative techniques were used to review the same studies, Hyde (1981) found that although differences were too small to be used as a primary explanation for females' lower levels of participation and productivity, males scored approximately one-fourth standard deviation higher than females.

Among the studies reviewed by Maccoby and Jacklin were 46 studies published in the five-year period prior to 1973 which compared males' and females' performance on measures of conservation, transitivity, seriation, and class inclusion. By and large, gender differences were consistently absent, despite the fact that most of the studies included in the sample were conducted with students at a developmental stage, adolescence, in which gender differences are often augmented. This finding was not anticipated, in view of the large number of earlier reports of male superiority on related varieties of tasks. Moreover, in a recent quantitative synthesis of studies on formal operational

thought, it was concluded that males outperform females on Piagetian tasks (Meehan, 1984).

For a number of reasons, the question of whether or not males perform at a higher level than females on cognitive tasks is an important one. The question is not only important for theoretical reasons, it is also important for educational practice. Piaget's theory is the most complete and systematic theory of cognitive development appearing in the literature (Flavell, 1963; Goldschmid, 1967; Hunt, 1961; Tuddenham, 1966), and as such it is the theory most often employed in designing instruction in science and mathematics. Moreover, Piagetian theory focuses on cognitive processes that many theorists consider critical to achievement in science. The finding that males and females consistently differ in their performance on tasks which presumably reflect these processes would have profound implications.

Important as this issue may be, there is currently no integrative review which provides a summary of the state of knowledge on the topic for an extended time frame and for a comprehensive body of literature. The broad goal of the present study was to contribute to the body of knowledge on gender differences in cognitive functioning by integrating the findings of existing research. The present synthesis complements the work of Meehan (1984) by examining more precisely the role of mediating variables known to affect cognitive performance and suspected of affecting the magnitude and direction of gender

differences, and by comparing studies conducted in America with those conducted in other countries.

The best methodology for the purposes of this study is meta-analysis (Glass, 1976). The meta-analytic approach involves transforming the findings of individual studies to a common metric, coding relevant characteristics of studies, and then using conventional statistical procedures to determine overall effects, subgroup effects, and relationships among characteristics of the studies and the findings. A special advantage of meta-analysis is that outcome indices are processed on an interval scale which renders them amenable to inferential statistical procedures.

The purpose of this study, then, was to conduct a meta-analysis of studies comparing the performance of male and female students on tests of cognitive functioning. Several questions guided the investigation:

1. Is there, overall, a significant difference between the performances of males and females on tests of cognitive functioning?
2. When differences occur, in what domains of cognitive performance, and in which operational schemata are male-female differences manifest? How large are the differences?
3. Are gender differences related to age? At what stage of intellectual development do they first become apparent?

4. Do the magnitude and direction of gender-related differences vary with the chronological time of investigation (e.g., are reported gender differences smaller in recent studies?)

Studies addressing these questions contain information on a melange of ancillary variables, such as test format, geographic region, and type of community. The precipitating roles played by many of these variables will be investigated in this study. Methodological deficiencies in the available research literature will receive special focus and be examined in relation to outcomes of the synthesis.

Effect Sizes and Vote Scores

Effect Sizes. Meta-analysis involves the expression of study outcomes in a common measure or standard term which permits statistical treatment. One commonly used measure is called the effect size: it represents the standardized difference between two comparison groups in performance on a dependent measure. In the present study, effect size was defined as the difference between male and female mean scores on a cognitive test, divided by the standard deviation for males. Preliminary analyses showed that while the mean standard deviations for males and females were essentially identical, the standard deviation associated with males' scores was slightly less variable. Since an effect size with the smaller variance is preferred--because it yields more consistent estimates across samples--the standard deviation associated with male scores was chosen for inclusion in the formula expressed as:

$$ES = \frac{\bar{X}m - \bar{X}f}{Sm}$$

where $\bar{X}m$ and $\bar{X}f$ represent male and female group means, respectively, and Sm represents the standard deviation of males' scores. In the case of studies that did not report means and and standard deviations for calculating effect sizes, formulas suggested by Glass (1977, pp. 267-270) were applied to compute them algebraically from other statistics. The effect size will be positive in studies yielding findings in favor of the male group and negative in studies in which findings are in favor of the female group.

Prior to all analyses of the data, a correction factor developed by Hedges (1981) was applied to each effect size in the study. The correction was utilized to correct for bias which Hedges found to be introduced by inflation of estimates associated with small sample sizes.

Vote Scores. Since many studies reported insufficient statistical information for the calculation of effect sizes by any of the formulas, the "vote score" method was employed. A report in which males outperformed females was labeled "M," a report in which females outperformed males was labeled "F," and a report indicating no gender differences was labeled "N." The tallies for the three categories were then totaled and compared. The vote score method complements the effect size procedure but is not adequate by itself because it yields no information concerning the magnitude of differences.

Literature Search, Selection, and Coding

To locate studies reporting separate scores for males and females on cognitive tasks, the science educational and psychological literatures were searched, beginning with 1981 and working backward to 1965. This period of time is sufficiently broad to allow analysis of changes in the magnitude and direction of gender differences across time.

Searching a large number of sources not only ensures the location of many studies but also ensures that biases in one source will be counteracted by those in other sources (Cooper, 1982). Thus, the first step in the literature search was a "cross-reference" computer search of library indexes to identify the appropriate data bases for further search. The preliminary search indicated that intensive computer searches of two library data bases were appropriate: Psychological Abstracts (PSYC) and Educational Resources Information Center (ERIC). In addition to the two computerized library searches, volumes of 14 major journals[3] were individually perused--page by page--for the years 1965 through 1981. This time-consuming procedure was considered essential for locating articles reporting gender differences because such differences, even if investigated, were not always mentioned in the abstracts. Reference sections of the located studies were also scanned for relevant titles.

The following criteria were used as guides for the selection of reports to be included in the synthesis:

1) The study contained a measure of cognitive functioning or logical performance (e.g., conservation, spatial ability, and reasoning);

2) the subject's performance was measured and reported in the English language;

3) the study provided data on students, kindergarten through early-college; and

4) summary statistics necessary for the calculation of effect sizes or vote scores were reported.

The search procedures located 70 reports which met the criteria for inclusion in the present synthesis. The reports included 60 journal articles, 9 dissertations and 1 published ERIC document. In all, 100 effect sizes were computed and 262 vote scores were documented. Complete references for studies selected for inclusion are provided in Appendix A, and a more detailed description of the studies appears in the first author's dissertation (Tohidi, 1982).

This review is not truly exhaustive of the literature on gender differences in cognitive functioning because although a large number of studies of cognitive functioning were located, many of them failed to meet one or more of the criteria for inclusion in the study. For example, many available studies did not investigate gender differences or did not report statistics necessary for calcuation of an effect size or vote score. Also, studies included in an earlier review conducted by Maccoby and Jacklin (1974) and synthesized by Hyde (1981) were excluded.

In order to characterize the studies and to express in quantitative terms each research context potentially mediating the results of the synthesis, a numerical coding scheme was developed. It includes key variables highlighted in the literature and reflects questions posed at the outset of the project. The coding categories, while being modified throughout the coding process, were ultimately clustered into four groups of variables: characteristics of the study, characteristics of the sample, characteristics of the instruments, and characteristics of the school(s) from which the sample was selected (Table 1).

Insert Table 1 about here

Categorization of Cognitive Tasks

Since data were not extensive enough to examine each cognitive task individually, the tasks were categorized into five cognitive domains: Conservation, Spatial Visualization, Cognitive Style, Reasoning, and Logical Thinking. Thus, an effect size for one cognitive domain, such as Conservation was composed of performance scores on a number of different tasks. Studies providing data for the cognitive domains and specific tasks are listed in Appendix B. In most cases, categorization of the tasks within the cognitive domains was guided by the discussion appearing in separate papers. Definitions used for confirming the categorizations are described in the following paragraphs.

Conservation. In Piaget's theory, the "schema of conservation" is viewed as a basic operational structure representing the ability of the child to perceive an invariant property of objects in spite of changes that take place in other properties (Za'rour, 1971). Conservation--of weight, length, volume, area--is considered one of the most basic constructs in the child's cognitive transition from the "preoperational" stage to "concrete operations" (Piaget, 1952; Piaget, Inhelder, & Szeminska, 1960).

Spatial Visualization. Spatial visualization or spatial conceptualization may include many abilities and elements of cognitive functioning, such as auditory localization, accurate maintenance of size-distance constancy, and tactual recognition of objects as they change orientation in space (Maccoby & Jacklin, 1974). Nevertheless, spatial visualization has been defined as "comprehension of imaginary movement in three-dimensional space" (Nash, 1979, p. 281), or as a "skill in identifying the appearance of an object whose parts have been turned in 3-dimensional space. Spatial visualization is closely related to other spatial factors such as spatial relations, spatial decontextualization or 'field dependence,' and spatial orientation" (Sherman & Fennema, 1978, p. 158).

Cognitive Style. Cognitive style is defined as "an habitual pattern or preferred strategy of information processing" (Cronbach & Snow, 1977, p. 375). Many researchers have identified two cognitive styles or "cognitive approaches," or

"ways of knowledge," or "types of intelligence." Witkin calls them "global versus analytical" styles (Witkin, 1979; Witkin, Dyk, Faterson, Goodenough, & Karp, 1962).

Reasoning. The category labeled "reasoning" includes tasks in which the subjects are required to generate hypotheses to explain a phenomenon, and then, employing both inductive and deductive reasoning, use evidence to confirm or contradict the hypotheses (Maccoby & Jacklin, 1974). According to Inhelder and Piaget (1958), "formal reasoning consists of a set of cognitive capabilities which enable an individual to deal systematically with abstractions and multivariate situations" (Bart, Baxter & Frey, 1980, p. 191). This category, therefore, includes studies examining Piagetian operational schemata of formal thought such as combinatorial operations, proportions, probability and correlation.

Logical Thinking. This cognitive domain includes studies that examine logical operations without specifying the type of mental operations. The term "logical thinking" indicates a general level of cognitive development of the subjects. It includes performance scores on tests dealing with things, words, thoughts, propositions, concepts, hypotheses, ideas and ideologies (Dulit, 1972).

Evaluating the methodological quality of the located studies was one of the most difficult aspects of the coding process. Attempts were made to code sample size, sample selection, bias in measurement or sampling, generalizability of results to other units or settings, validity, and reliability of the measures. In

many cases, however, information necessary for the evaluation of the quality of the studies was not reported. For example, many studies failed to provide an adequate description of the sample, or failed to report on validity and reliability of the measures involved. Information concerning the control of factors potentially jeopardizing internal and external validity of the study (Campbell & Stanley, 1967) was frequently lacking.

When coding of the 70 studies was complete, 10 randomly selected studies (15% of the sample) were coded by a second coder, in order to establish inter-coder reliability. There was 92% agreement on the coded variables. Agreement on quantitative elements of the studies was almost perfect while agreement concerning subjective evidence was less pronounced.

Statistical Analyses

A general linear models approach to analysis of variance (Tatsuoka, 1975), which controls for intercorrelations among variables, was used to analyze the data. Effect sizes and vote scores served as dependent variables in all analyses. In the case of age and publication date, which were distributed as continuous variables, linear regression analysis was utilized. Chi-square analyses were applied to vote scores. Follow-up analyses were undertaken with subsets of the data in an attempt to understand more fully those characteristics of the study, the sample, the measuring instruments, and the school which contributed to the magnitude of gender differences.

File Drawer Effect

The File Drawer Effect (Greenwald, 1975) refers to bias introduced by the greater tendency of significant or expected results to be published in the literature. However, formulas for calculating the hypothetical number of studies necessary to invalidate results of published studies (Rosenthal, 1979) require statistics which, because the investigation of gender differences was often not among the primary purposes of the investigations, were not reported in many of the studies included in this meta-analysis. By the same token, as noted by Meehan (1984), the fact that gender differences were not among the primary hypotheses of the study obviates the necessity of an investigation of publication bias associated with such differences.

Units of Analysis

Endemic to meta-analysis is the problem of dependence among findings from the same samples of subjects, a phenomenon which leads to inflated Type I error rates. Most combinatorial methods are based on the assumption that "study" rather than "result" is the unit of analysis and that each effect size is thus independent of the others. Yet it is often the case that "study" cannot be considered the unit of analysis without aggregating findings above the level at which many interesting relationships can be studied. A number of solutions for this situation have been suggested in the literature, but interpretive difficulties are associated with each of them. The common procedures of randomly selecting only one effect size from each study (Meehan, 1984), or pooling all individual effect sizes within a single

study and calculating an average effect size, both result in the loss of important data. The application of the Jackknife technique (Mosteller & Tukey, 1968) also requires that data be discarded. Because the effect sizes from the studies being synthesized tended not to yield unsatisfactory standard errors for the mean effect sizes (cf. Glass, 1978), and because preliminary combining of effect sizes within studies or including only randomly selected effect sizes would have made it impossible to investigate differences by task, it was felt that the disadvantages of an a priori correction outweighed the advantages. Thus, since only seven studies providing effect sizes included scores from the same subjects on more than one cognitive task, the effect size for the individual gender comparison was used as the unit of analysis. This fact is reiterated where relevant to interpretation of the findings.

Results and Discussion

Direction and Magnitude of Gender Differences

As can be seen in Table 2, gender differences in cognitive functioning tend to favor males. Although the effect sizes are significantly different from zero (confidence coefficient = .95), their values tend to be low. The largest effect size, reported for foreign countries, is a positive .56, which means that in studies conducted outside the U.S., males' performance scores on cognitive tests represented in the sample are, on the average, a little more than a half standard deviation higher than females'

scores. In studies conducted within the United States, the effect size is a positive .27, indicating that males' scores are approximately one-fourth of a standard deviation higher than females'. Of the 262 vote scores, 52% reflected a male advantage and 10% indicated a female advantage.

Insert Table 2 about here

Differences of this overall magnitude are not consistent with the notion that cognitive ability per se is the exclusive or primary reason why females are underrepresented in science classes and technical professions. While not large, the differences are pervasive, however, and suggest more detailed analyses. The specific domains in which gender differences occur and specific variables which affect the magnitude of these differences are of special interest to those concerned with teaching mathematics and science to both boys and girls. We turn now to an examination of these issues.

Variables Affecting the Magnitude of Gender Differences

As can be seen in Table 3, a number of categorical and continuous variables tend to have an effect ($p < .10$) on the magnitude of effect sizes in cognitive functioning. These will be discussed in the order of their appearance in the tables and figures.

Insert Table 3 about here

Cognitive Domain

Studies contributing data to each of the cognitive domains and task categories are listed in Appendix B.

Conservation. Results in Table 3 show that Cognitive Domain has a significant effect on the variance of effect sizes. It can be seen that within the subcategories of this variable, the largest effect size (1.02) is associated with the schema of conservation. It should be noted, however, that the effect size on conservation is produced here by three independent subgroups, ages 7, 8, and 9, of a single research project conducted by Za'rour (1971) on a Lebanese sample (n = 132) and dealt with conservation of weight. The Chi-square analysis of vote scores associated with Cognitive Domain also shows a similar male advantage in the performance of conservation tasks.

In spite of questions regarding the stability of the finding of gender differences in conservation, possible reasons for relatively large effect sizes and vote scores in favor of males should be considered. Does the male role allow for greater experience in organizing, manipulating and measuring than does the female role? Gender-typed patterns of play activities (Carpenter & Huston-Stein, 1980) that provide boys with greater freedom of outside exploration, manipulation, and perceiving of objects before and after different transformations (Goldschmid, 1967) may result in boys' greater exposure to materials or experiences which are mechanical or science-related in nature. Increased exposure may also stem from a generally greater sociocultural opportunity which urges boys toward cognitive

achievement. For example, in the study of conservation in Lebanon, Za'rour (1971) attributes male superiority to the Lebanese culture, in which "boys are more often sent to buy things from the neighborhood store or from a peddler and are instructed to check the accuracy of the weighing" (p. 393). In a meta-analysis of conformity research, Cooper (1979) provided evidence that girls are more apt than boys to conform, particularly in face-to-face situations. This tendency limits girls' experience with cognitively challenging materials and ideas.

Spatial ability. The effect size (.37) for spatial ability reflects data from 37 studies of visual-spatial ability, as can be seen in Table 3. This effect size is much smaller than that associated with conservation and parallels results reported earlier in quantitative integrations of studies appearing in the literature. For example, Hyde (1981) applied meta-analytic techniques to the studies of gender difference cited by Maccoby and Jacklin (1974). She reported means for males and females on visual-spatial tasks which were less than a half standard deviation apart, as can be seen in Figures 1a and 1b where the results reported by Hyde can be compared with those reported here. The effect size revealed in the present study is still smaller than the one reported by Hyde, suggesting, in accordance with interpretations of Sherman (1978) and Hyde, that even for the supposedly "well-established" gender differences in spatial ability (Maccoby & Jacklin, 1974; Liben, 1978; Signorella &

Jamison, 1978), the difference is very small: it accounts for less than 5% of the population variance.

Insert Figures 1a and 1b about here

Despite its small magnitude, the very pervasiveness of a male advantage in reports on spatial visualization has highlighted a number of practical implications (Boles, 1980; Eliot & Fralley, 1976; Hyde, 1981; Petersen, 1979; Sherman & Fennema, 1978; Treagust, 1980) and generated discussion and research on the origins of gender differences. Some contend that biological mechanisms are involved (Benbow & Stanley, 1980; Bock & Kolakowski, 1973; Hartlage, 1970; Stafford, 1972). Others have stressed environmental/experiental, sociocultural, motivational, and measurement factors in explaining the observed gender differences in spatial ability and other areas of cognitive functioning (Berry, 1966; Cohen, 1980; Connor & Serbin, 1977; Dwyer, 1979; Farmer, 1978; Fennema & Sherman, 1978; Nash, 1979; Sherman, 1967; Smedsland, 1977; Steinkamp, 1984). A third view maintains that gender differences occur through a complex interaction between social and biological factors (Harris, 1978; Wittig & Petersen, 1979).

At this point, a definitive conclusion cannot be drawn concerning the etiology of gender differences in spatial ability. The environmental/experiential theories have been most successful in providing convincing evidence in support of their arguments, but they have not demonstrated the specific mechanisms whereby the

environmental and sociocultural factors operate to generate gender differences.

Cognitive style. As can be seen in Table 3, an effect size of .44 resulted from five studies of cognitive style. Sherman (1967) argued that gender differences in performance on some cognitive tests--for example, the Embedded Figures Test--are produced by the spatial component of the tests rather than by the analytic component. In other words, she maintains that since these tests load heavily on a spatial factor, "females' poorer performance on visual field independence tasks may simply be a reflection of undeveloped visual-spatial abilities and have nothing to do with decontextualization" (quoted by Maccoby & Jacklin, 1974, p. 97).

Maccoby and Jacklin (1974) report that both visual-spatial tasks that involve disembedding and those that do not show a similar male superiority in early adolescence. According to their review there is no indication that the disembedding process contributes to gender differences beyond what would be produced by the visual-spatial component of the tasks.

Studies reported by Feather (1968), Lipton and Overton (1971), Maccoby (1966), Werdelin (1961), and Witkin (1971; 1979) support the notion that the genders perform similiarly on measures of field-dependency that involve disembedding (e.g., auditory spatial location, haptic pattern tracing, body steadiness, verbal-embeddedness anagrams) but not visual-spatial perception. Piburn (1977) suggested that "sex and field-dependence appear to operate quite independently in terms of their influence on

analytical problem-solving ability. In particular, they seem to differentially affect different schemata in the Piagetian model of formal thought" (p. 14).

Given the small magnitude of effect sizes obtained in this synthesis, and considering the questions and ambiguity of earlier findings, no definitive answer can be provided concerning gender differences in cognitive style. Further research on this issue is needed.

Reasoning and Logical Thinking. Although Piaget did not discuss gender differences, it has been suggested (Piburn 1977) that the relatively lower scores of girls in reasoning at the formal level leads them to become increasingly unsuccessful and disinterested in science at the secondary and college levels. The point that can be made here is that existing gender differences in reasoning and logical thinking are too small (effect size = .27 and .18) to be used as a primary explanation for females' low level of participation in technical courses and careers.

In order to gain a deeper understanding of patterns of gender differences among the various cognitive domains, foreign and domestic samples were examined separately, as shown in Table 4 for the specific tasks (Studies contributing data to the cognitive domains and task categories are listed in Appendix B). In general, effect sizes in foreign samples are larger than those in the United States and Canada. On several tasks, females scored higher than males. A slightly negative effect size (-.03), indicating a female advantage, emerged in the United

States' sample for "combination of different tasks." With respect to "classification and seriation," there is a small negative effect size in the U.S. sample (ES = -.17) and a much smaller negative effect size in the foreign sample (ES = - .03). In terms of vote scores from 17 studies on classification and seriation , one favored males, three favored females, and 13 showed no gender differences. Although the magnitude is small, the stronger performance of females on these tasks merits further consideration.

Insert Table 4 about here

Bridgham (1969) attributes the superiority of girls in the performance of classification and seriation tasks to the technically-oriented occupations of the fathers in his sample. He argues that "an interest in science is considered a 'normal' thing for a boy; for a girl it may be 'normal' only if there is a family influence making it so" (p. 126). He also suggests that the effects of other variables such as interest, locus of attention, and positive attitudes might have confounded the results and produced the gender differences emerging in the sample.

Females' slightly higher scores in the area of classification are of special interest, in light of related observations and research results. In the first place, classification of living organisms is universally taught in kindergarten and elementary school and in introductory courses in biological science in early junior high, before students are

fully aware that science has been classified by society as a male domain (Cowan, 1971; Weinreich-Haste, 1978). This curricular pattern is consistent with findings of Steinkamp and Maehr (1983; 1984) who report that girls' attitudes are more positive than boys' in school biology, chemistry, and botany. It is also consistent with numerous reports (Cole, 1979; Comber & Keeves, 1973) that the proportion of females in the biological sciences is higher than the proportion in the physical sciences. Taken together, these findings suggest a pattern of considerable reliability and constitute an example of an early predisposition interacting with environmental factors to provide a predictable pattern of choice and achievement.

Response mode

As anticipated, the results show that the format of the task provides a context in which the relationship between gender and cognitive performance varies. When response mode (oral, written, or manipulative), and response format (free or structured), were examined, significant patterns emerged. The superiority of boys over girls in cognitive performance tends to increase when the testing requires active manipulation of apparatus (Table 3) and, in the foreign sample, with a free-response mode, as shown in Figure 2 where foreign and domestic data are compared.

Insert Figure 2 about here

In a study on gender differences in problem-solving ability, Graybill (1975) explained that with test instruments requiring manipulation, boys generally appeared more confident in handling the equipment and more sure of their movements; girls asked permission to perform operations more than boys did, and were less free than boys in their pursuit of the solutions to the problems presented. In contrast, girls' performances exceeded those of boys when the experiment required a minimum of manipulation and the kinds of skills required were similar to those required for the performance of household tasks.

These findings suggest that the format of cognitive tasks may be discriminatory against female subjects (Dale, 1970; Lawson, 1975). Jamison and Signorella (1980) have argued that in the case of some cognitive measures such as the Water-Level Task, "the apparatus tends to create the impression of a laboratory exercise. It is not that the instructions for the task are difficult to understand or carry out, but rather that the task and concept convey an aura of science. If this is true, then the Water-Level Task may be subtly biased against females who have been socialized to have less confidence in their abilities to deal with things scientific" (p. 346).

These findings supplement patterns emerging in earlier work when aspects of test content were examined (Dwyer, 1979; Graybill, 1975; Jamison & Signorella, 1980; Kagan, 1964; Lawson & Shepherd, 1979; Linn, 1981; Milton, 1959; & Sherman, 1978). As early as 1959, Milton found that when adolescents were presented with mathematics problems involving typically male activities.

males consistently obtained higher scores than females. However, when the problems, identical in terms of logic and computation, had a "feminine content" such as cooking or gardening, the performance of females improved remarkably.

Region of country. The disparity between boys' and girls' scores on cognitive tasks is largest in the western regions of the United States. The reasons for this are not clear. Whether or not the pattern reflects a more traditional psychosocial environment in the Western states remains a matter for speculation.

Sample Selection

Sample selection emerged as a significant variable ($p = .08$), with random selection of subjects associated with larger effect sizes. Sampling has been considered an important methodological issue in a number of recent literature reviews (Sherman, 1978). Hyde (1981) argues that careful attention has not been paid to sampling and sample size in prior research or reviews of gender-related differences in cognition. One problem, for example, is that a large sample size increases the chance of finding a significant gender difference, even when the difference is actually very small and accounts for a small percentage of the variance.

Year of publication

The significant effect resulting when effect size was regressed on year of publication ($F = 9.95$; $p < .002$; $df = 1,99$) shows a gradual decrease in the magnitude of gender differences across chronological time of investigation (Figure 3). This

pattern confirms the results of earlier studies which report smaller gender differences in more recent studies (Maccoby & Jacklin, 1974; Petersen, 1979). The trend implies that the overall estimate of effect size (mean effect size = .32), which is based upon studies conducted since 1965, may now be actually too large to reliably reflect the most current situation. If gender differences have their origin in psychosocial factors, decreases in the magnitude of gender differences would be more compatible with recent societal changes in sex roles than would increases. Another possibility is that the smaller effect sizes in recent years merely reflect improvement in test construction and more reliable measurement of cognitive tasks. The data neither support nor refute these notions.

Insert Figure 3 about here

Age

It is of interest that age did not significantly affect gender differences in the studies included in this synthesis. This finding was not anticipated, in view of evidence for the cumulative nature of sex-role effects across age (Nash, 1979). Recent attempts to broaden the sex-role concepts of boys and girls may serve as a partial explanation for cross-age consistency found in this study.

As noted, Chi-square analyses of vote scores yielded results which tend to complement the patterns reported for effect sizes.

For types of community, however, the pattern of vote scores differs from the pattern of effect sizes. In terms of effect sizes, the male advantage is of similar magnitude in urban, suburban, and rural communities (Table 3), whereas in terms of vote scores, the males' score superiority is less pronounced in suburban communities (Chi-square = 22.28; df = 6; $p < .01$). In those communities, neither gender scored higher in 72% of the cases.

Insert Table 5 about here

Summary and Conclusions

The primary purpose of this synthesis was to determine the magnitude and direction of gender-related differences in performance on cognitive tasks reported in studies published during the time period 1965 through 1981. In the 70 studies synthesized here, it was found that gender-related differences do occur but that they are small. Boys slightly outperform girls in the cognitive domains of conservation, spatial visualization, cognitive style, reasoning, and logical thinking, whereas girls perform better than boys in specific tasks involving classification and seriation.

While differences of this magnitude cannot be used as an explanation for females' underrepresentation in science-related and technical careers, the very pervasiveness of the male advantage is of great interest to those concerned with theoretical questions. How can the small yet pervasive

differences be explained? Does this synthesis--which examined all objectively-measured data appearing in the studies which could conceivably impinge on performance disparities between the sexes--provide information which can be construed as support for either the genetic or the psychosocial explanation of gender differences?

While the pervasiveness of the difference across what are assumed to be basic cognitive abilities appears to suggest that a genetic-origins-hypothesis cannot be ignored, there is little positive evidence that gender differences are of genetic origin. First, if the finding that gender differences are smaller in recent years is replicable and not an artifact of psychometric refinements, it would provide negative support for a genetic-origins-hypothesis. If hereditary endowment were the major explanation, the size of the gender differences would not have shown significant variation during the 15-year time period spanned in this synthesis. Second, gender differences are larger in cultures and regions where males are traditionally nurtured in less restrictive environments than females (cf. Za'rour, 1971) and are exposed to independent and explorative activities during what Piaget calls "the concrete stage of development." Such variations in boys' and girls' activities are noted by Sutton-Smith (1961) in an early cross-cultural study of children's games.

Perhaps the strongest evidence for the psychosocial explanation of gender differences is provided by Piaget himself. Indeed, Piaget's theory is based on the notion that performance

on cognitive tasks is malleable and related to the richness and variety of the child's early experiences with the environment. The very experiences which Piaget deemed essential to the development of strong cognitive abilities--measuring and comparing, estimating time and distance, and problem-solving in the context of play--are the experiences to which, in many cultures, males more than females are exposed in the early years (cf. Carpenter & Huston-Stein, 1980; Lever, 1978; Lewis, 1972; Vandenberg, 1978; Za'rour, 1971). This line of argument suggests that early cognitive socialization may be a critical factor in creating the small but pervasive differences revealed in the present analysis (cf. also, Steinkamp & Maehr, 1983).

All in all, the research evidence to date clearly indicates that males and females differ in certain cognitive abilities which are presumably important to science achievement. The differences are small, but they are pervasive. The research evidence does not specify the origins of these differences with any high degree of certainty. While the nature, pervasiveness and consistency of the differences may suggest a role for genetics there is also evidence that psychosocial factors are important. Certainly, it would be foolish to conclude that gender differences in aptitude for school science are inevitable and not malleable.

References

Armstrong, J. M. (1979). A national assessment of achievement and participation of women in mathematics. Final Report to the National Institute of Education on Grant No. NIE-G-77-0061. Denver, CO: Education Commission of the States.

Astin, H. S. (1969). The woman doctorate in America. New York: Russell Sage Foundation.

Bart, W. M., Baxter, J., & Frey, S. (1980). The relationship of spatial ability and sex to formal reasoning capability. Journal of Psychology, 104, 191-198.

Benbow, C., & Stanley, J. (1980). Sex differences in mathematical ability: Fact or artifact? Science, 210, 1262-1264.

Berry, J. W. (1966). Temne and Eskimo perceptual skills. International Journal of Psychology, 1, 207-229.

Bock, D. R., & Kolakowski, D. (1973). Further evidence of sex-linked major gene influence on human spatial visualizing ability. American Journal of Human Genetics, 25, 1-14.

Boles, D. B. (1980). X-linkage of spatial ability: A critical review. Child Development, 51, 625-635.

Bridgham, R. G. (1969). Classification, seriation, and the learning of electrostatics. Journal of Research in Science Teaching, 6, 118-127.

Campbell, D. T., & Stanley, J. C. (1967). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Carpenter, C. J., & Huston-Stein, A. (1980). Activity structure and sex-typed behavior in preschool children. Child Development, 51, 862-872.

Cohen, R. F., & Cohen, M. R. (1980). Opening new doors: Taking sex-role stereotyping out of science and mathematics. School Science and Mathematics, 705, 566-572.

Cole, J. R. (1979). Fair science: Women in the scientific community. New York: Free Press.

Comber, L. C., & Keeves, J. P. (1973). Science education in nineteen countries. New York: Wiley.

Connor, J. M., & Serbin, L. A. (1977). Behaviorally based masculine-and-feminine-activity-preference scales for preschoolers: Correlates with other classroom behaviors and cognitive tests. Child Development, 48, 1411-1416.

Cooper, H. M. (1979). Statistically combining independent studies: A meta-analysis of sex differences in conformity research. Journal of Personality and Social Psychology, 37(1), 131-146.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. Review of Educational Research, 52(2), 291-302.

Cowan, R. (1971). School children's perception of a preference for school subjects. Unpublished Master's thesis, Edinburgh.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods. New York: Irvington.

Dale, L. G. (1970). The growth of systematic thinking: Replication and analysis of Piaget's first chemical experiment. Australian Journal of Psychology, 22, 277-286.

Dulit, E. (1972). Adolescent thinking a la Piaget. Journal of Youth and Adolescence, 1, 281-301.

Dwyer, C. A. (1979). The role of tests and their construction in producing apparent sex-related differences. In M. A. Wittig & A. C. Petersen (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press.

Eliot, J., & Fralley, J. S. (1976). Sex differences in spatial abilities. Young Children, 31, 487-498.

Farmer, H. (1978). What inhibits achievement and career motivation in women? The Counseling Psychologist, 6, 12-15.

Feather, N. T. (1968). Changes in confidence following success or failure as a predictor of subsequent performance. Journal of Personality and Social Psychology, 9, 38-46.

Fennema, E., & Sherman, J. (1978). Sex-related differences in mathematics achievement and related factors: A future study. Journal for Research in Mathematics Education, 9, 189-203.

Flavell, J. H. (1963). The developmental psychology of Jean Piaget. Princeton, NJ: Van Nostrand.

Fogelman, K. R. (1972). Piagetian tests and sex differences--II. Educational Researcher, 12, 154-155.

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G. (1978). Integrating findings: The meta-analysis of research. In L. Shulman (Ed.). Review of research in education, Vol. 5. Itasca, IL: Peacock.

Goldschmid, M. L. (1967). Different types of conservation and non-conservation and their relation to age, sex, IQ, MA, and vocabulary. Child Development, 38, 1229-1246.

Graybill, L. (1975). Sex differences in problem-solving ability. Journal of Research in Science Teaching, 12, 341-346.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. Psychological Bulletin, 82(1), 1-20.

Harris, L. J. (1978). Sex differences in spatial ability: Possible environmental, genetic, and neurological factors. In M. Kinshourne (Ed.), Asymmetrical function of the brain. London: Cambridge University.

Hartlage, L. C. (1970). Sex-linked inheritance of spatial ability. Perceptual and Motor Skills, 31, 610.

Hedges, L. V. (1981). Distribution theory for Glass' estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128.

Hedges, L. V., & Olkin, I. (1983). Clustering estimates of effect magnitude from independent studies. Psychological Bulletin, 93(3), 563-573.

Hunt, J. V. (1961). Intelligence and experience. New York: Ronald.

Husen, T. (1967). International study of achievement in mathematics. Stockholm: Almqvist & Wiksell.

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using w^2 and d. American Psychologist, 36, 892-901.

Jamison, W., & Signorella, M. (1980). Sex-typing and spatial ability: The association between masculinity and success on Piaget's water-level task. Sex-Roles, 6, 345-353.

Kagan, J. (1964). Acquisition and significance of sex typing and sex-role identity. In M. L. Hoffman & L. W. Hoffman (Eds.), Review of child development research. New York: Russell Sage Foundation.

Keeves, J. (1973). Differences between the sexes in mathematics and science courses. International Review of Education, 19, 47-62.

Lawson, A. E. (1975). Sex differences in concrete and formal reasoning ability as measured by manipulative tasks and written tasks. Science Education, 59, 397-405.

Lawson, A. E. & Shepherd, G. D. (1979). Syntactical complexity in written language and cognitive development at the formal level. Science Educaton, 63, 73-81.

Lever, J. (1978). Sex differences in the complexity of children's play and games. American Sociological Review, 43, 471-483.

Lewis, M. (1981). Culture and gender roles: There is no unisex in the nursery. Psychology Today, 5, 54-57.

Liben, L. S. (1978). Performance on Piagetian spatial tasks as a function of sex, field dependence and training. Merrill-Palmer Quarterly, 24, 97-110.

Linn, M. C. (1981, April). Male-female differences in formal thought. Paper presented at Society for Research in Child Development, Boston.

Lipton, C., & Overton, W. F. (1971). Anticipatory imagery and modified anagram solution: A developmental study. Child Development, 42, 615-623.

Maccoby, E. E. (1966). Sex differences in intellectual functioning. In E. E. Maccoby (Ed.), The development of sex differences. Stanford: Stanford University Press.

Maccoby, E. E., & Jacklin C. (1974). The psychology of sex differences. Stanford, CA: Stanford University Press.

Meehan, M. A. (1984). A meta-analysis of sex differences in formal operational thought. Child Development, 55, 1110-1124.

Milton, G. A. (1959). Sex differences in problem solving as a function of role appropriateness of the problem content. Psychological Reports, 5, 705-708.

Mosteller, F. M., & Tukey, J. W. (1968). Data analysis, including statistics. In G. Lindzey and A. Aronson (Eds.), Handbook of social psychology (2nd Ed.). Reading, MA: Addison-Wesley.

Nash, S. C. (1979). Sex role as a mediator of intellectual functioning. In M. A. Wittig & A. C. Petersen (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press.

Petersen, A. C. (1982, April). Biological correlates of spatial ability and mathematical performance. Paper presented at the Annual Meeting of American Educational Research Association, New York.

Piaget, J. (1952). The child's conception of number. New York: Humanities.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). The child's conception of geometry. New York: Basic Books.

Piburn, M. D. (1977, March). Sex, field dependence and formal thought. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching. Cincinnati, OH.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. Psychological Bulletin, 83(3), 638-641.

Schonberger, A. K. (1976). The interrelationship of sex, visual spatial abilities and mathematical problem solving. Unpublished doctoral dissertation, University of Wisconsin.

Sherman, J. A. (1967). A problem of sex differences in space perception and aspects of of intellectual functioning. Psychological Review, 74, 290-299.

Sherman, J. A. (1978). Sex-related cognitive differences: An essay on theory and evidence. Springfield, IL: C. C. Thomas.

Sherman, J. A., & Fennema, E. (1978). Distribution of spatial visualization and mathematical problem solving scores: A test of Stafford's x-linked hypothesis. Psychology of Women Quarterly, 3, 157-167.

Signorella, M. L., & Jamison, W. (1978). Sex differences in correlations among field-dependence, spatial ability, sex role orientation, and performance on Piaget's water-level task. Developmental Psychology, 14, 689-690.

Smedslund, J. (1977). Piaget's psychology in practice. British Journal of Educational Psychology, 47, 1-6.

Stafford, R. (1972). Hereditary and environmental components of quantitative reasoning. Review of Educational Research, 42, 183-201.

Steele, L., & Wise, L. K. (1979, April). Origins of sex differences in high school mathematics achievement and participation. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Steinkamp, M. (1984). Motivational style as a mediator of adult achievement in science. In M. W. Steinkamp & M. L. Maehr (Eds.). Advances in motivation and achievement: Women in science. Greenwich, CT: JAI Press.

Steinkamp, M., Harnisch, D., Walberg, H. J., & Tsai, S. (1985). Cross-national gender differences in mathematics attitude and achievement among thirteen-year-olds. The Journal of Mathematical Behavior, 4(3).

Steinkamp, M. & Maehr, M. L. (1983). Affect, ability, and science achievement: A quantitative synthesis of correlational research. Review of Educational Research, 53, 369-396.

Steinkamp, M., & Maehr, M. L. (1984). Gender differences in motivational orientations toward achievement in school science: A quantitative synthesis. American Educational Research Journal, 21(1), 39-59.

Sutton-Smith, B. (1961). Cross-culture study of children's games. Philadelphia: Yearbook American Philosophical Society, 426-429.

Tatsuoka, M. (1975). The general linear model, a "new" trend in analysis of variance. Selected topics in advanced statistics, No. 7. Champaign, IL: Institute for Personality and Ability Testing.

Tohidi, N. (1982). Sex differences in cognitive performance on Piaget-like tasks: A meta-analysis of findings. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Treagust, D. F. (1980). Gender-related differences of adolescents in spatial representational thought. Journal of Research in Science Teaching, 17, 91-97.

Tuddenham, R. D. (1966). Jean Piaget and the world of the child. American Psychologist, 21, 207-217.

Vandenberg, B. (1978). Play and development from an ethological perspective. American Psychologist, 33, 724-738.

Werdelin, I. (1961). Geometrical ability and the space factors in boys and girls. Sweden: University of Lund.

Weinreich-Haste, H. (1978). Sex differences in fear of success among British students. British Journal of Social and Clinical Psychology, 17, 37-42.

Witkin, H. A. (1979). Socialization, culture and ecology in the development of group and sex differences in cognitive style. Human Development, 22, 358-372.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. (1962). Psychological differentiation. New York: Wiley.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). A manual for the Embedded Figures Test. Palo Alto, CA: Consulting Psychologists.

Wittig, M. A., & Petersen, A. C. (1979). Sex-related differences in cognitive functioning. New York: Academic Press.

Za'rour, G. I., & Panouri-Kilaniotis, A. L. (1971). Analytic styles of categorization and conservation abilities among Greek Cypriot children. Science Education, 61, 47-56.

Footnotes

[1] The research on which this paper is based was supported by NSF-RISE Grant No. 80-07857, Martin L. Maehr, Principal Investigator.

[2] This paper is based on a doctoral dissertation conducted at the University of Illinois by the first author under the supervision of the second and third authors.

[3] The following journals were searched: Science Education; Journal of Research in Science Teaching; Developmental Psychology; Child Development; Psychology of Women Quarterly; Sex Roles; Human Development; School Science and Mathematics; Journal of Psychology; British Journal of Psychology; Australian Journal of Psychology; Merrill-Palmer Quarterly; Educational Research; and Science and Children.

Table 1

Coding Variables
Characteristics of Study
Year of publication
Source of reference
Quality of journal
Sample selection
Sample size ... Number of males
... Number of females
Type of analysis
Direction of results (favoring males/females/neither)
Effect size (the criterion or dependent variable)
Characteristics of Sample
Country of sample
Region of country
Ethnicity
Type of community
Socioeconomic status
Achievement level
Age

Table 1 (continued)

Characteristics of Instrument

- Type of reliability
- Reliability index
- Type of validity
- Source of test
- Administration (individual/group)
- Stimulus mode
- Response mode
- Free/structured response
- Cognitive domain
- Task or schemata

Characteristics of School

- Public/private
- Code/separate
- Instruction type

Table 2

Means and Standard Deviations of Effect Sizes (Total Sample)

Studies	Mean Effect Size	Standard Deviation of Effect Sizes	Number of Effect Sizes
United States and Canada	.27*	.39	81
Foreign	.56*	.77	19
Total	.32*	.50	100

*95% Confidence Interval did not include zero.

Table 3

Results of One-Way Analysis of Variance on the Significant Variables (Total Sample)

Categorical Variables	Mean Effect Size	Number of Effect Sizes
Cognitive Domain $F = 2.21$, $df = 4/99$, $p = .07$		
Conservation	1.02	3
Spatial ability	.37	37
Cognitive style	.44	5
Reasoning	.27	38
Logical thinking	.18	17
Response Mode $F = 2.35$, $df = 3/99$, $p = .08$		
Oral	- .07	3
Written	.21	33
Manipulative	.35	47
Mixed or not specified	.54	17
Region of Country $F = 2.40$, $df = 3/99$, $p = .07$		
North, Northeast, Northwest	.19	26
South and Southwest	.22	29
West	.45	20
Not specified	.48	25
Sample Selection $F = 2.32$, $df = 3/99$, $p = .08$		
Sample poorly specified	- .08	3
Individual subjects randomly selected	.48	33
Subset of entire school or classroom	.14	3
One or more entire school(s) or classroom(s)	.26	61

Table 3 (continued)

Type of Community $F = 2.83$, $df = 3/99$, $p = .04$

Urban	.46	20
Suburban	.46	18
Rural	.48	11
Mixed of not specified	.19	51

ible 4

esults of One-Way Analyses of Variance (Type of Task)[a]

Type of Task	U. S. and Canadian Sample		Foreign Sample		Total Sample	
	Mean Effect Size	Number of Effect Sizes	Mean Effect Size	Number of Effect Sizes	Mean Effect Size	Number of Effect Sizes
1. Horizontality/Verticality (Water-level task)	.68	1	-	-	.68	1
2. Geometric, topological tasks	.19	7	.65	2	.30	9
3. Embedded Figures Test (Field dependence/independence)	.39	4	-	-	.39	4
4. Classification and Seriation	-.17	3	-.03	4	-.09	7
5. Proportional	.40	7	-	-	.40	7
6. Combination and Permutation	.13	19	.65	10	.31	29
7. Probability and Correlation	.19	4	-	-	.19	4
8. Propositional and Abstraction	.43	1	-	-	.43	1
9. Separation and Controlling Variables	.80	1	-	-	.80	1
10. Conservation	-	-	1.02	3	1.02	3
11. Author-developed tasks	.39	29	-	-	.39	29
12. Combination of different tasks	-.03	5	-	-	-.03	5
	F=1.83, df=11/80, p=.07		F=1.24, df=3/18, p=.33		F=1.53, df=11/99, p=.14	

[a] See Appendix B for studies providing data for the separate tasks.

Figure Captions

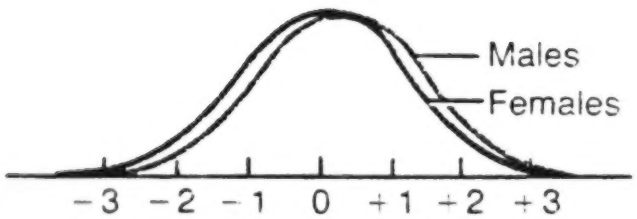
Figure 1a. Two normal distributions with means .50 SD apart, that is, with an effect size of .50. This is approximately the magnitude of the gender differences in visual-spatial ability and field articulation reported by Hyde (1981)

Figure 1b. Two normal distributions with means .37 SD apart, that is, with an effect size of .37. This is approximately the magnitude of the gender differences in visual-spatial ability obtained by the present synthesis.

Figure 2. Relationship between reponse format and magnitude of gender differences.

Figure 3. Relationship between year of publication of studies and magnitude of gender differences.

A.


Males
Females
−3 −2 −1 0 +1 +2 +3
Mean Effect Size

B.

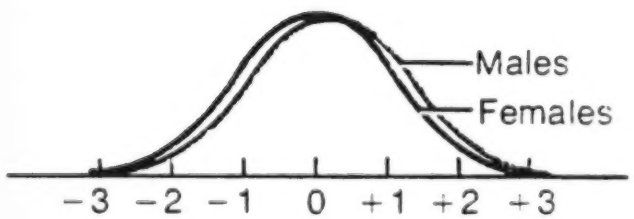

Males
Females
−3 −2 −1 0 +1 +2 +3
Mean Effect Size

Figure 1.

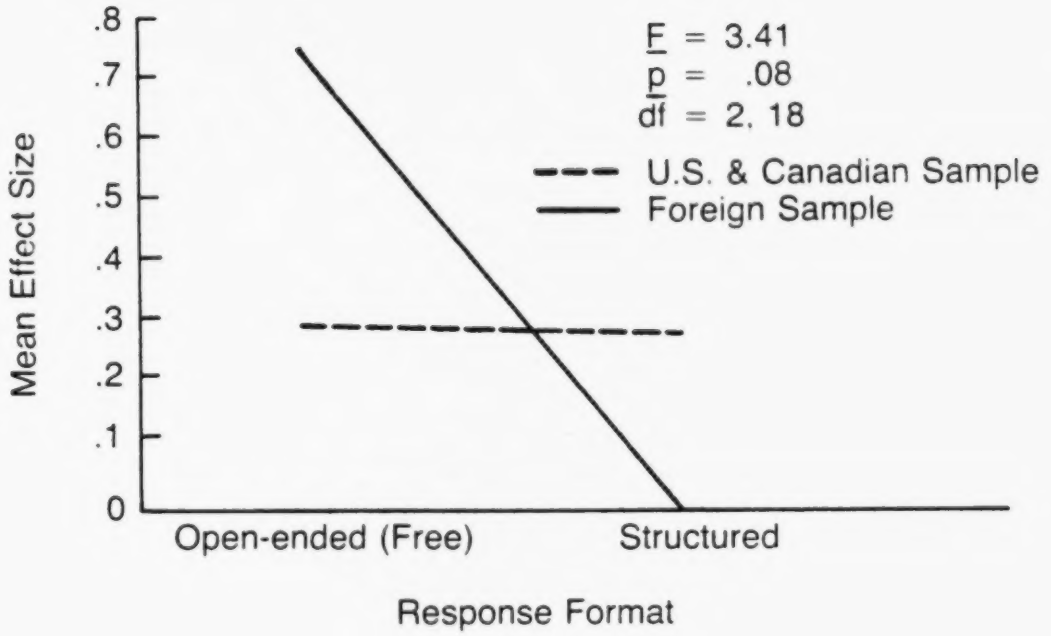

.8
.7
.6
.5
.4
.3
.2
.1
0
Mean Effect Size
F = 3.41
p = .08
df = 2, 18
U.S. & Canadian Sample
Foreign Sample
Open-ended (Free)
Structured
Response Format

Figure 2.

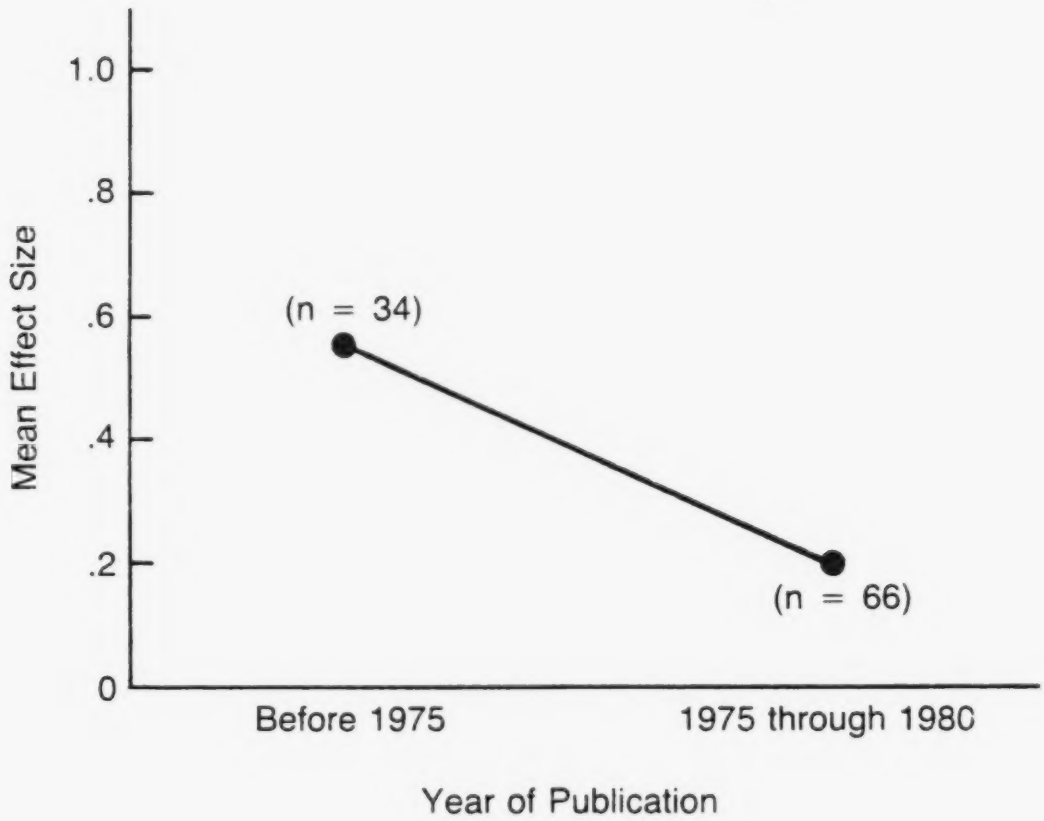

1.0
.8
.6
.4
.2
0
Mean Effect Size
(n = 34)
(n = 66)
Before 1975
1975 through 1980
Year of Publication

Figure 3.

Appendix A

List of Synthesized Studies

Anderson, R. D. Children's ability to formulate mental models to explain natural phenomena. Journal of Research in Science Teaching, 1965, 3, 326-332.

Bart, W. M., Baxter, J., & Frey, S. The relationship of spatial ability and sex to formal reasoning capability. Journal of Psychology, 1980, 104, 191-198.

Bridgham, R. G. Classification, seriation, and the learning of electrostatics. Journal of Research in Science Teaching, 1969, 6, 118-127.

Carlson, G. R. Location of a point in Euclidean space by children in grades one through six. Journal of Research in Science Teaching, 1976, 13, 331-337.

Cleminson, R. W. A cognitive study of three fifth-grade classrooms on selected Piaget-type tasks dealing with science-related concepts. Unpublished doctoral dissertation, The University of Iowa, 1970.

Cohen, H. G. The scaling of six topological Piagetian groupings, as well as the effect that certain selected variables have on the attainment of these groupings and some of their homologs in the logical domain. Journal of Research in Science Teaching, 1978, 15, 115-125.

Dale, L. G. The growth of systematic thinking: Replication and analysis of Piaget's first chemical experiment. Australian Journal of Psychology, 1970, 22, 277-286.

DeLuca, F. P. Measurement of logical thinking: An electronic equivalent of Piaget's first clinical experiment. Journal of Research in Science Teaching, 1977, 14, 539-544.

DeLuca, F. P. Application and analysis of an electronic equivalent of Piaget's first chemical experiment. Journal of Research in Science Teaching, 1979, 16, 1-11.

DeLuca, F. P. Application of cluster analysis to the study of Piagetian stages of intellectual development. Journal of Research in Science Teaching, 1981, 18(1), 51-59.

Dettrick, G. W. A study of the structure of Piagetian logical and infralogical groupings within the concrete operational period of cognitive development. Dissertation Abstracts International, 1975, 7719-A.

Douglas, J. D., & Wong, A. C. Formal operations: Age and sex differences in Chinese and American children. Child Development, 1977, 48, 689-692.

Doyle, Jerry J. The order of attainment of eight projective groupings: An analysis of Piaget's spatial model. Journal of Research in Science Teaching, 1980, 17(1), 55-58.

Dulit, E. Adolescent thinking a la Piaget. Journal of Youth and Adolescence, 1972, 1, 281-301.

Fennema, E., & Sherman, J. Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal, 1977, 14, 51-71.

Field, T. W., & Cropley, A. J. Cognitive style and science achievement. Journal of Research in Science Teaching, 1969, 6, 2-10.

Flanery, R. C., & Balling, J. D. Developmental change in hemispheric specialization for tactile spatial ability. Developmental Psychology, 1979, 15(4), 364-372.

Fogelman, K. R. "Piagetian tests and sex differences--II." Educational Research, 1972, 12(2), 154-155.

Friot, F. E. The relationship between an inquiry teaching approach and intellectual development. Unpublished doctoral dissertation, University of Oklahoma at Norman, 1970.

Gillespie, J. P. A study of concept development and communication abilities in young children using variations in leaf morphology. Journal of Research in Science Teaching, 1972, 9, 57-64.

Goldschmid, M. L., "Different types of conservation and nonconservation and their relation to age, sex, IQ, MA, and vocabulary," Child Development, 1967, 38, 1229-1246.

Graybill, Letitia. Sex differences in problem-solving ability. Journal of Research in Science Teaching, 1975, 12(4), 341-346.

Gross, George R. A study of the influence of formal high school science and math course participation, sex, and IQ on level of cognitive development as measured by standard Piagetian and alternate tasks. Dissertation Abstracts International, 1978, 38(7-A), 4071.

Hademenos, J. G. A comparative study of Piaget-type conservation tasks. School Science and Mathematics, 1974, 74, 680-686.

Hammond, J., & Raven, R. The effects of a structured learning sequence on the achievement on compensatory tasks. Journal of Research in Science Teaching, 1973, 10, 257-262.

Hensley, J. H. An investigation of proportional thinking in children from grades six through twelve. Unpublished doctoral dissertation, The University of Iowa, 1974.

Jahoda, G. Sex and ethnic differences on a spatial-perceptual task: Some hypotheses tested. British Journal of Psychology, 1980, 71, 425-431.

Jahoda, G. On the nature of difficulties in spatial-perceptual tasks: Ethnic and sex differences. British Journal of Psychology, 1979, 70, 351-363.

Jamison, W., & Signorella, M. Sex-typing and spatial ability: The association between masculinity and success on Piaget's water-level task. Sex Roles, 1980, 6(3), 345-353.

Karplus, R., Karplus, E., Formisano, M., & Paulsen, A. C. A survey of proportional reasoning and control of variables in seven countries. Journal of Research in Science Teaching, 1977, 14, 411-417.

Keating, D., & Schaefer, R. Ability and sex differences in the acquisition of formal operations. Developmental Psychology 1975, 11, 531-532.

Keating, D. P., & Clark, L. V. Development of physical and social reasoning in adolescence. Developmental Psychology, 1980, 16(1), 23-30.

Kelly, G. N., & Kelly, J. T. Basic physical concepts: Do college students have them? Journal of College Science Teaching, 1978, 8(1), 29-31.

Kelly, J. T., & Kelly, G. N. Science activity centers: Basic concepts of the physical world. Science and Children, 1978, 16(3), 25-26.

Keough, B. K., & Keough, J. F. Pattern copying under three conditions of an expanded spatial field. Developmental Psychology, 1971, 4, 25-31.

Killiam, C. R. Cognitive development of college freshmen. Journal of Research in Science Teaching, 1979, 16(4), 347-350.

Kishta, M. A. Proportional and combinational reasoning in two cultures. Journal of Research in Science Teaching, 1979, 16, 439-443.

Kishta, M. A. Cognitive levels and linguistic abilities of elementary school children. Journal of Research in Science Teaching, 1979, 16, 67-71.

Lawson, A. E., & Shepherd, G. D. Syntactical complexity in written language and cognitive development at the formal level. Science Education, 1979, 63(1), 73-81.

Lawson, A. E. Sex differences in concrete and formal reasoning ability as measured by manipulative tasks and written tasks. Science Education, 1975, 59(3), 397-405.

Leskow, S., & Smock, C. Developmental changes in problem-solving strategies: Permutations. Developmental Psychology, 1970, 2(3), 412-422.

Liben, L. S. Performance on Piagetian spatial tasks as a function of sex, field dependence and training. Merrill-Palmer Quarterly, 1978, 24, 97-110.

Linn, M. C., & Levine, D. I. Adolescent reasoning: Influence of question format and type of variables on ability to control variables. Science Education, 1978, 62, 377-388.

Lowery, L. F., & Allen, L. R. Socio-economic status and sex differences in visual resemblance sorting tasks at the first grade level. Journal of Research in Science Teaching, 1970, 7, 57-65.

Maxwell, J. W., Croake, J. W., & Biddle, A. P. Sex differences in the comprehension of spatial orientation. The Journal of Psychology, 1975, 91, 127-131.

Nelson, G. K. Concomitant effects of visual, motor, and verbal experiences in young children's concept development. Journal of Educational Psychology, 1976, 68(4), 466-473.

Norman, D. K. A comparison of children's spatial reasoning: Rural Appalachia, suburbans, and urban New England. Child Development, 1080, 51, 288-291.

O'Brien, T., & Shapiro, B. The development of logical thinking in children. American Education Research Journal, 1968, 5, 531-542.

Odegaard, G. E. An investigation of Piaget's groupings: Seriation and projective space. Dissertation Abstracts International, 1975, 7959-A.

Ott, W. R. An investigation of certain of Piaget's infralogical groupings dealing with projective and Euclidean space. Dissertation Abstracts International, 1979, 39(8-A), 4848-4849.

Piburn, Michael D. "Sex, field dependence and formal thought." Paper presented at the annual meeting of the National Association for Research in Science Teaching, Cincinnati, Ohio, March 22-24, 1977. ED 139 635.

Popp, L., & Raven, R. The effects of response format of a structured learning sequence on third grade children's classification achievement. Journal of Research in Science Teaching, 1972, 9, 177-184.

Raven & Adrian. Relationships among science achievement, self-concept, and Piaget's operative comprehension. Science Education, 1978, 62, 471-479.

Reesink, C. J. The performance of secondary school students on Piaget-type tasks concerning projective and Euclidean space. Dissertation Abstracts International, 1977,

Robertson, W. W., & Richardson, E. The development of some physical science concepts in secondary school students. Journal of Research in Science Teaching, 1975, 12, 319, 329.

Ryman, D. Teaching methods, intelligence, and gender factors in pupil achievement on classification task. Journal of Research in Science Teaching, 1977, 14, 401-409.

Saarni, C. I. "Piagetian operations and field independence as factors in children's problem-solving performance," Child Development, 1973, 44, 338-345.

Shayer, M., & Wylam, H. Distribution of Piagetian stages of thinking in British middle and secondary-school children. British Journal of Educational Psychology, 1978, 48, 62-70.

Sherman, J. A., & Fennema, E. Distribution of spatial visualization and mathematical problem solving scores: A test of Stafford's x-linked hypotheses. Psychology of Women Quarterly, Winter 1978, 3(2), 157-167.

Siegler, R. S., & Richards, D. D. Development of time, speed and distance concepts. Developmental Psychology, 1979, 15(3), 288-298.

Signorella, M. L., & Jamison, W. Sex differences in correlations among field-dependence, spatial ability, sex role orientation, and performance on Piaget's water-level task. Developmental Psychology, 1978, 14, 689-690.

Smith, W. S., & Schroeder, C. K. Instruction of fourth grade girls and boys on spatial visualization. Science Education, 1979, 63, 61-66.

Stone, M. The development of the intersituational generality of formal thought. Unpublished doctoral dissertation, University of Illinois at Urbana, 1966.

Thomas, H., & Jamison, W. On the acquisition of understanding that still water is horizontal. Merrill-Palmer Quarterly, 1975, 21(1), 31-44.

Treagust, D. F. Gender-related differences of adolescents in spatial representational thought. Journal of Research in Science Teaching, 1980, 17, 91-97.

Uzgiris, I. C. "Situational generality of conservation," Child Development, 1964, 35, 831-841.

Vasta, R., Regan, K. G., & Kerley, J. Sex differences in pattern copying: Spatial cues or motor skills? Child Development, 1980, 51(3), 932-934.

Willemsen, E., & Raynolds, B. Sex differences in adults' judgment of the horizontal. Developmental Psychology, 1973, 8, 309.

Za'rour, G. I. Conservation of weight across different materials by Lebanese school children in Beirut. Science Education, 1971, 55, 387-394.

Za'rour, G. I., & Panaouri-Kilaniotis, A. L. Analytic styles of categorization and conservation abilities among Greek Cypriot children. Science Education, 1977, 61, 47-56.

Appendix B

Studies Providing Data for the Cognitive Domains and Specific Tasks

Cognitive Domain[a]	Reference[b]
Conservation	Saarni (1973); Goldschmid (1967); Hademenos (1974); Shayer (1978); Robertson (1975); Fogelman (1972); Lawson (1975); Kishta (1979); Za'rour (1977); Cleminson (1970).
Spatial Visualization	
Horizontality/Verticality Water-level Task	Signorella (1978); Thomas (1975); Maxwell (1975); Kelly (1978); Treagust (1980); Jamison (1980); Liben (1978); Willemson (1973); Fennema (1977).
Geometric, Topological Tasks	Smith (1979); Cohen (1978); Sherman (1978); Fennema (1977).
Cognitive Style	
Embedded Figures	Saarni (1973); Signorella (1978); Liben (1978).
Field Dependence/Independence	
Reasoning (Concept Mastery)	
Classification and Seriation	Ryman (1977); Lowery (1970); Popp (1972); Bridgham (1969); Kishta (1979); Gillespie (1972); Dettrick (1974); Odegaard (1975).
Proportional Reasoning	Lawson (1975); Keating (1980); Dulit (1972); Kishta (1979); Gross (1977); Piburn (1977); Bart (1980); Cleminson (1970); Hensley (1974).
Combination and Permutation	Saarni (1973); Dale (1970); Deluca (1977, 1979, 1981); Kishta (1979); Piburn (1977); Dulit (1972); Leskow (1970).

Appendix B (continued)

Cognitive Domain[a]	Reference[b]
Probability and Correlation	Piburn (1977); Bart (1980).
Propositional and Abstraction	Karplus (1977); Hammond (1973); Stone (1966); Friot (1970).
Separation and Controlling Variables	Shayer (1978); Linn (1978); Lawson (1975); Keating (1980); Bart (1980); Deluca (1981).
Logical Thinking	
Author-developed Tasks	O'Brien (1968); Keogh (1971); Za'rour (1971); Bridgham (1969); Field (1969); Anderson (1965); Vasta (1980).
Combinations of Different Tasks	Keating (1975); Killian (1979); Piburn (1977); Za'rour (1977); Raven (1978); Siegler (1979); Nelson (1976); Uzgiris (1964); Graybill (1975); Douglas (1977); Lawson (1979); Goldschmid (1967).

[a] Several of these studies reported insufficient statistical information to calculate effect sizes but were coded for vote scores.

[b] Complete references appear in Appendix A.